
Chapter 14

Enhancing Access to UK Renal Registry Data through Innovative Online Data Visualisations

Afzal Chaudhry^{a,b}, Terry Feest^b

^aCambridge University NHS Hospitals Foundation Trust, Cambridge, UK; ^bUK Renal Registry, Bristol, UK

Summary

- It is now possible to interactively interrogate, analyse and produce displays of UK Renal Registry data online at http://www.renalreg.com/n_portal/pages/main/registryportal.php. This enables any individual to interrogate the validated anonymised UKRR datasets.
- This bespoke web-based interactive data portal provides a focussed point of access to a variety of graphical display formats and analyses of UKRR data.
- Centre-specific reports can be produced – including a colour-coded dashboard summary as well as both funnel plots and longitudinal statistical process control charts for a range of clinical parameters.
- Interactive flash-based longitudinal Statistical Process Control charts are available on a per-centre and per-parameter basis allowing for a more detailed review of performance over time.
- There are Rosling/Gapminder-style motion charts on a per-parameter basis simultaneously detailing performance and activity data from multiple centres interactively over time (more details below).
- There is an interactive graphical pivot chart solution using OLAP (online analytical processing) technology allowing users to design and export their own charts/analyses in real-time using UKRR data.
- The portal will empower the UK renal community in the comparative analysis of delivered renal care ultimately hopefully leading to enhanced quality improvement over time.

Background

The UK Renal Registry (UKRR) has grown from strength-to-strength since its inception in 1997 and now receives data from every renal centre in England, Wales and Northern Ireland with a more limited dataset being contributed annually by the Scottish Renal Registry. Each year the report provides detailed comparisons of the activity and performance of each of the submitting centres across a range of clinical parameters. Whilst this data and its associated analyses are valued by a wide range of stakeholders (the renal centres themselves, the NHS Renal healthcare Commissioners and allied patient-related groups e.g. the National Kidney Federation), it is recognised that as the volume of data and analyses grows with time, so does the need to present this increasingly complex information in an accessible and clinically informative manner which is responsive to, and reflects the nature of, the enquiries made by those seeking to access the data. In particular, the need to reflect changes in activity and performance over time (longitudinal data) as opposed to the predominantly ‘snapshot’ cross-sectional data contained in the annual report is essential if the UKRR is to achieve

its goal of monitoring renal care in the UK thus leading to improvements in the quality and efficiency of this very same care.

Online Interactive Geographical Maps

It was in this context that in 2009 the UKRR launched the world’s first interactive maps (figure 14.1 and table 14.1) detailing the achievement of quality measures in the care of dialysis patients spanning a five-year period. Initially these maps were confined to a variety of haematological and biochemical parameters (table 14.2) based around healthcare commissioning geographies. They were soon followed by the addition of national public health datasets (cardiovascular mortality, indices of social deprivation and ethnic distributions courtesy of the East Midlands Public Health Observatory, table 14.2) based on these same geographies allowing for a limited exploration of correlations between these parameters and UKRR data. Renal centre-based maps displaying UKRR data alone are now also available (table 14.2).

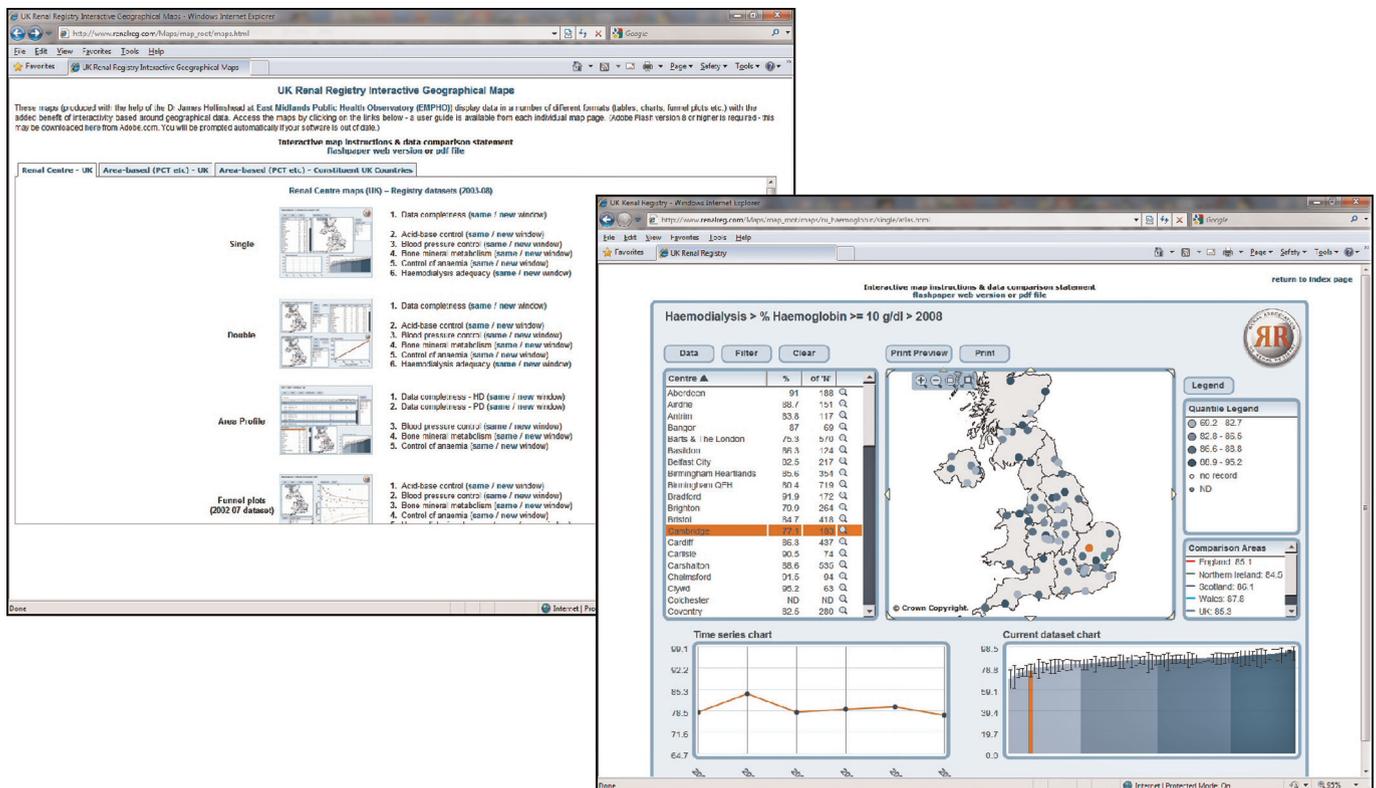


Fig. 14.1. Interactive Geographical Maps – http://www.renalreg.com/Maps/map_root/maps.html

Table 14.1. UK Renal Registry online interactive map types

Geography basis	Type	Description
Area-based Centre-based	Single	View containing a thematic map, table and chart indicating spatial/geographical patterns and temporal trends for a specific indicator of interest.
Area-based Centre-based	Double	View containing two thematic maps allowing for the comparison of patterns and relationships between two indicators for the same geography type.
Area-based Centre-based	Area Profile	View allowing users to select a geographical 'area' and to chart a cross-section of selected key indicators. This provides an 'at-a-glance' assessment of the selected 'area' in comparison with other 'areas', including regional and/or national geography regions.
Centre-based	Funnel Plot	View of cross-sectional data containing a thematic map, table and chart indicating spatial/geographical patterns as well as a funnel plot with upper and lower 95% and 99.9% confidence intervals.

The Statistical Comparison of Data

Historically many of the graphs in the various UKRR reports describing, for example, the proportion of patents in a given centre achieving a given biochemical standard, have been presented in the form of a

'caterpillar' plot with centres listed on the x-axis in order of the percentage achievement. At the same time it was recognised that these types of plots perhaps encouraged inappropriate statistical comparisons between centres. In the context of the online maps, the guidance states:

Table 14.2. UK Renal Registry online interactive map datasets

Country	Geography	UKRR data	Public Health Data
England	Primary Care Trust	RRT incidence and prevalence rates Bone mineral metabolism (Ca, PO ₄ , PTH) Control of anaemia (Hb, ferritin) Haemodialysis adequacy	Circulatory disease mortality 2005–2007 ^a Index of Multiple Deprivation 2007 ^b Ethnicity – Office of National Statistics 2006 ^c
Northern Ireland	Local District	RRT incidence and prevalence rates Bone mineral metabolism (Ca, PO ₄ , PTH) Control of anaemia (Hb, ferritin) Haemodialysis adequacy	Circulatory disease mortality 2005–2007 ^d Northern Ireland Multiple Deprivation Index 2005 ^d Ethnicity – 2001 Census ^c
Scotland	Health Board	RRT incidence and prevalence rates	Circulatory disease mortality 2006–2008 ^e Scottish Index of Multiple Deprivation 2009 ^e Ethnicity – 2001 Census ^c
Wales	Local Health Board	RRT incidence and prevalence rates Bone mineral metabolism (Ca, PO ₄ , PTH) Control of anaemia (Hb, ferritin) Haemodialysis adequacy	Circulatory disease mortality 2002–2004 ^f Welsh Index of Multiple Deprivation 2008 ^g Ethnicity – 2001 Census ^c
All	Renal Centre	As maps above, plus (where data is available): Data 'completeness' Acid-base control (Bicarbonate) Blood pressure control	No data aggregated by renal centre available

Data Sources:

^a Compendium of Clinical and Health Indicators

^b National Statistics, Department of Communities and Local Government

^c Office for National Statistics

^d The Northern Ireland Statistics and Research Agency

^e ScotPHO

^f National Public Health Service for Wales

^g Welsh Assembly Government

‘Caution should be used when comparing data between many areas using maps such as these. Unless you have a priori comparisons in mind and make only those specific comparisons then there will be problems.

...The funnel plots allow for the identification of renal centres which fall outside of the upper or lower 95% and 99.9% confidence intervals. These compare all centres with the average rather than being centre to centre comparisons like you are tempted to make from the ‘caterpillar’ style plots. You are still making multiple comparisons (comparing each of about 70 centres with the average) but using the 99.9% intervals instead of the 95% ones makes some adjustment for this.’

More recently therefore the UKRR has made increased use of funnel plots for presenting such data in the annual report.

Contemporaneously with the introduction of these maps, Dr Alex Hodsman, Renal Registry Research Fellow, was advancing the application of additional robust and well-validated statistical analytical techniques to UKRR data. One area of focus in her research was the use of statistical process control charts, originally developed by Dr Walter Shewart for the Bell Telephone Company in 1924 (see: http://en.wikipedia.org/wiki/Walter_A._Shewart).

Dr Shewart realised that variation in data can occur for a number of reasons. Assuming that data collection and processing is robust then there are principally two types of variation: ‘common cause’ and ‘special cause’ variation. All processes demonstrate some degree of random variation, which is known as ‘common cause’ variation and a particular process is said to be ‘in control’/‘stable’ if it demonstrates only ‘common cause’ variation. However, unexpected events or situations arising in a process can result in ‘special cause’ variation. In such circumstances, a process is said to be ‘out of control’/‘unstable’. Variation of this type needs to be identified and, if genuine, investigated further.

Statistical process control (SPC) charts can be used to differentiate between these two types of variation (using a variety of ‘rules’) and can be plotted as either cross-sectional or longitudinal charts using either static (a snapshot in time) or dynamic (a series of data points over time) data respectively. Funnel plots are a means of displaying cross-sectional/static data whilst longitudinal run charts can be used to display longitudinal/dynamic data (see: <http://www.indicators.scot.nhs.uk/SPC/Main.html>).

The UK Renal Registry Online Data Portal

Following on from:

- considerable positive feedback regarding the interactive maps,
- a desire to extend the statistical rigor applied to UKRR data (as outlined above), and
- feedback from many centres about their own accessibility to UKRR data with increasingly frequent requests for timely access

the UKRR now wishes to report the extension of this online interactive strategy to the deployment of a bespoke interactive data portal (Summer 2011). The aim of the portal is to provide a focussed point of access to a variety of graphical display formats and analyses of UKRR data, with figures 14.2A, B, C and D showing the different data presentation options:

- A. Centre-specific reports – a distillation of annual UKRR data including a colour-coded dashboard summary as well as both funnel plots and longitudinal statistical process control charts for a range of clinical parameters. The dashboard describes for each clinical parameter:
 - the centre’s performance (percentage achievement of standard) for that year
 - a numerical and colour-coded comparison with the previous year
 - whether or not the centre is an outlier on a funnel plot in that year
 - whether or not the centre’s performance exhibits ‘special cause’ variation over time on a longitudinal SPC
 - the mean percentage achievement of all centres in the same region for that year
 - a colour coded median rank (with 95% confidence intervals) amongst all centres for that parameter based upon a standard statistical simulation model.
- B. Interactive flash-based longitudinal SPC charts on a per-centre and per-parameter basis allowing for a more detailed review of performance over time. These charts are the interactive correlates of those available in the centre-specific reports.
- C. Rosling/Gapminder-style motion charts on a per-parameter basis simultaneously detailing performance and activity data from multiple centres interactively over time (more details below).

OLAP Pivot Charts

An OLAP (online analytical processing) cube is a data structure that allows for the rapid analysis of data categorised by a number of *dimensions*. Each of the elements of a dimension can be described using a hierarchy which is a series of parent–child relationships in a generational type structure. So, each parent is derived by the aggregation of its child members and in turn, parent members at any one level may be further aggregated as the children of another parent at the next higher level. As an example, September 2003 can be aggregated into Quarter 3 2003, which in turn can be aggregated into Year 2003 – all of these members belong to the ‘Time’ dimension.

The numerical data in the cube goes to form the *measures* of the cube. These are usually summary calculations (e.g. minimum, maximum, mean, total etc.) for the various data points aggregated across the dimensions of the cube at the level of the hierarchy selected (e.g. the mean percentage attainment of a given standard for all centres in Quarter 3 2003).

One of the strengths of an OLAP cube is the ability to drill-up and to drill-down through the hierarchical levels of one or more dimensions (e.g. starting by looking at data aggregated by year, then drilling down to data aggregated by quarter and so on). Another is the ability to rapidly re-orientate the cube so as to look at the same data but from an entirely different perspective e.g. instead of looking at data through the ‘Time’ dimension, the data could be viewed from the perspective of the ‘Location’ dimension which might start with the four Home countries of the UK, drilling down to a regional level and then down to the level of the individual renal centres themselves. In reality, data is usually explored and aggregated simultaneously across multiple dimensions (e.g. geographical region, time, RRT modality etc.) – but the power of the OLAP cube means that changes in the selected dimensions/hierarchies are reflected in the displayed graphic within seconds rather than in the minutes or hours that might be needed if aggregating this data via a traditional relational database query.

In the implementation of an OLAP cube used in this data portal, the OLAP processing is done in the Flash/Adobe Flex front end client running on the local computer following transfer of all the raw data from a MySQL database on the UKRR web server. A more

efficient strategy would be for the OLAP processing to take place on the UKRR’s server following submission of a query and for only the results of the query to be passed to the requesting client. This is to be considered in a future development.

At present the power of the OLAP cube is tied to an interactive pivot chart style of front end where the user may ‘drag and drop’ parameters and in addition, may select individual members of different hierarchies in order to generate their own personalised charts from validated UKRR datasets. These charts may then be exported in either a jpg or pdf format for use locally.

Future Plans and Summary

All of the online interactive visualisations use the same validated UKRR datasets as those used in each respective year’s annual report. Currently (Summer 2011), the online maps house data from the 2003–2008 UKRR datasets (except for the funnel plots which house the 2002–2007 datasets). The centre-specific reports in the online portal are derived from the 2007 dataset whilst the SPC charts, the motion charts, and the OLAP pivot chart all use data from the 2002–2007 datasets.

Over the next few months more of the UKRR’s annual datasets will be uploaded into portal. More complex comparisons will be developed to take advantage of the power of the motion charts and as the volume of data grows so the technical structure of the OLAP charting system will be revised to improve the responsiveness of the system as outlined above. In addition, the aim is to more fully integrate the geographical maps into the data portal itself. Further refinements will also be made in response to the feedback received from users of the portal.

We believe that this work – ‘Enhancing Access To Registry Data Through Innovative Online Data Visualisations’ – builds strongly on the wealth of information arising from the high-quality validated UKRR datasets, and that both the maps and the portal will empower and engage the UK renal community in the comparative analysis of delivered renal care ultimately leading to enhanced quality improvement over time.

Conflicts of interest: none